INFORMATION AND CONTROL 41, 156–164 (1979)

Mutual Information, Strong Equivalence, and Signal Sample Path Properties for Gaussian Processes[1]

CHARLES R. BAKER

Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514

Let (S_t) and (N_t), t in $[0, T]$ be stochastic processes with almost all paths in $\mathscr{L}_2[0, T]$, jointly Gaussian. Relations are obtained between the average mutual information of (S_t) and $(S_t + N_t)$, strong equivalence of the measures induced by $(S_t + N_t)$ and (N_t), and the almost sure sample path properties of (S_t).

INTRODUCTION

Average mutual information, absolute continuity of measures (non-singular detection), and sample path analysis of stochastic processes have been areas of considerable research interest for many years. In addition to results within these areas, relations between the areas are of interest. For example, in the problem of detecting a possibly non-Gaussian signal in independent Gaussian noise, it is possible to give sufficient conditions for non-singular detection purely in terms of the analytical characteristics of the signal sample paths, without knowing anything about the statistical behavior of the signal (Baker, 1973a).

In this paper, we obtain tight relations between these several properties when all the processes concerned are Gaussian. These results are extensions of previous results due to T. S. Pitcher (1963) (see also the survey by Osteyee and Good (1974)). Weaker relations for non-Gaussian signals are given elsewhere (Baker, 1978).

We assume throughout that (S_t) and (N_t), $t \in [0, T]$, are measurable zero-mean stochastic processes on (Ω, β, P), jointly Gaussian. Both processes are assumed to have almost all paths in $\mathscr{L}_2[0, T]$. Under these assumptions, (S_t), (N_t), and $(S_t + N_t)$ have covariance operators R_S, R_N, and R_{S+N}, respectively. For example, the covariance operator of (N_t) is the integral operator in $L_2[0, T]$ having the covariance function of (N_t) as its kernel. Further, each of the three processes induces a probability measure on the Borel sets of $L_2[0, T]$ via its path map. We denote these induced measures as μ_S, μ_N, and μ_{S+N}, (for (S_t),

[1] Research supported by Office of Naval Research Contract N00014-75C-0491.

0019-9958/79/050156-09$02.00/0

(N_t), $(S_t + N_t)$, respectively). For example, $\mu_N(A) = P\{\omega : N(\omega) \in A\}$, where $N(\omega)$ is the path of (N_t) (in $\mathscr{L}_2[0, T]$) at the point ω in Ω. The processes $(S_t + N_t)$ and (N_t) are non-singular (resp., singular) if the measures μ_{S+N} and μ_N are mutually absolutely continuous or *equivalent* (resp., not mutually absolutely continuous). If the two measures are mutually absolutely continuous, we write $\mu_{S+N} \sim \mu_N$, and if they are orthogonal, we write $\mu_{S+N} \perp \mu_N$. Since μ_{S+N} and μ_N are Gaussian, either $\mu_{S+N} \sim \mu_N$ or $\mu_{S+N} \perp \mu_N$ (Feldman, 1958; Hajek, 1958).

In order to discuss average mutual information (AMI), one must also consider joint measures on the Borel σ-field of $L_2[0, T] \times L_2[0, T]$. Thus $\mu_{S,N}$ is the joint measure induced by the path map of the pair of stochastic processes (S_t, N_t); similarly for $\mu_{S,S+N}$ and $(S_t, S_t + N_t)$. $\mu_S \otimes \mu_N$ is the usual product measure; $\mu_{S,N} = \mu_S \otimes \mu_N$ if (S_t) and (N_t) are mutually independent stochastic processes. We recall (Dobrushin, 1963) that if $\mu_{X,Y}$ is a joint measure and $\mu_X \otimes \mu_Y$ its (unique) product measure, then the average mutual information of $\mu_{X,Y}$ (which we denote as $I(X, Y)$) is infinite unless $\mu_{X,Y}$ is absolutely continuous with respect to $\mu_X \otimes \mu_Y$. When $\mu_{X,Y} \ll \mu_X \otimes \mu_Y$, then

$$I(X, Y) = \iint \log \left[\frac{d\mu_{X,Y}}{d\mu_X \otimes \mu_Y} (x, y) \right] d\mu_{X,Y}(x, y).$$

Now consider the detection problem. When the two measures are zero-mean and Gaussian, $\mu_{S+N} \sim \mu_N$ if and only if (see, e.g., Rao and Varadarajan, 1963) there exists a Hilbert–Schmidt operator T in $L_2[0, T]$ which does not have -1 as an eigenvalue and which satisfies $R_{S+N} = R_N^{1/2}(I + T)R_N^{1/2}$. If $\mu_{S+N} \sim \mu_N$ and both measures are zero-mean and Gaussian, and the operator T is also trace-class, then μ_{S+N} and μ_N are said to be *strongly* equivalent, denoted $\mu_{S+N} \sim^s \mu_N$ (Hajek, 1962).

The first major result in the area of this paper was obtained by Pitcher (1963), who proved the following. Suppose that (S_t) and (N_t) are independent Gaussian processes *and* that $\mu_{S+N} \sim \mu_N$. Then the following are equivalent: (a) $\mu_{S+N} \sim^s \mu_N$; (b) $I(S, S + N) < \infty$; (c) almost all paths of (S_t) belong to the range of $R_N^{1/2}$. In relating $I(S, S + N)$ and absolute continuity between μ_{S+N} and μ_N, it has been shown (Baker, 1970) that if (S_t) and (N_t) are jointly Gaussian, and $R_{S+N} \geqslant R_N$, then $I(S, S + N) < \infty$ implies $\mu_{S+N} \sim^s \mu_N$.

The relation of sample path properties to strong equivalence and finite AMI can be analyzed from the following result (Baker, 1973a); If (S_t) is Gaussian, then almost all paths of (S_t) belong to range $(R_N^{1/2})$ ($\mu_S[\text{range}(R_N^{1/2})] = 1$) if and only if $R_S = R_N^{1/2} T R_N^{1/2}$ for T a trace-class operator. Thus, if (S_t) and (N_t) are independent Gaussian processes, then $\mu_{S+N} \sim^s \mu_N$ if and only if almost all paths of (S_t) belong to range $(R_N^{1/2})$.

If $(S_t + N_t)$ is not Gaussian, then we say that $\mu_{S+N} \sim^s \mu_N$ provided that $\mu_{S+N} \sim \mu_N$ and that $R_{S+N} = R_N^{1/2}(I + T)R_N^{1/2}$ for a trace-class operator T. This agrees with the definition for $(S_t + N_t)$ Gaussian. It has been shown

(Baker, 1977) that if (S_t) is not Gaussian, and (S_t) and (N_t) are not independent, then the fact that $R_{S+N} = R_N^{1/2}(I + T)R_N^{1/2}$ for T trace-class, with -1 not an eigenvalue of T, does not imply that $\mu_{S+N} \sim \mu_N$. Relations between mutual information, strong equivalence, signal sample path properties, and the signal-to-noise ratio of a quadratic-linear test statistic have been obtained (Baker, 1978) for the case where (S_t) is allowed to be non-Gaussian. Of course, those relations are weaker than the relations obtained here.

It should perhaps be noted that the idea of strong equivalence is of particular interest when $(S_t + N_t)$ and (N_t) are each a segment of a stationary Gaussian process with rational spectral density. In this case, if Φ_{S+N} is the spectral density of $(S_t + N_t)$ and Φ_N is the spectral density of (N_t), then $S + N$ and N are strongly equivalent if and only if $\lim_{|\lambda|\to\infty} \Phi_{S+N}(\lambda)/\Phi_N(\lambda) = 1$. Moreover, if this is not satisfied, then μ_{S+N} and μ_N are orthogonal (Hajek, 1962).

Mutual Information, Strong Equivalence, and Sample Path Properties

We assume hereafter that (S_t) and (N_t) are jointly Gaussian, so that $\mu_{S,N}$ is a Gaussian measure. First, we require several lemmas.

Lemma 1 (Baker, 1973b). *Let $\mu_{X,Y}$ be a joint Gaussian measure on the Borel σ-field of $L_2[0, T] \times L_2[0, T]$, with projections μ_X and μ_Y. Let P_X (resp., P_Y) be the projection operator with range equal to* $\overline{\text{range}(R_X)}$ *(resp.,* $\overline{\text{range}(R_Y)}$*). $\mu_{X,Y}$ has a cross-covariance operator R_{XY} which has the decomposition $R_{XY} = R_X^{1/2}TR_Y^{1/2}$, $\| T \| \leqslant 1$, $T = TP_Y = P_XT$. Conversely, given μ_X and μ_Y as Gaussian measures on $L_2[0, T]$, any operator of the form $R_{XY} = R_X^{1/2}TR_Y^{1/2}$, $\| T \| \leqslant 1$, defines a joint Gaussian measure $\mu_{X,Y}$.*

If μ_{XY} is induced by jointly Gaussian processes (X_t, Y_t), then the cross-covariance operator of $\mu_{X,Y}$ can be represented by an integral operator having the cross-covariance function of (X_j) and (Y_t) as its kernel.

Lemma 2 (Baker, 1970). *Suppose $\mu_{X,Y}$ is a joint Gaussian measure on $L_2[0, T] \times L_2[0, T]$. Then $I(X, Y) = I(\mu_{X,Y})$ is finite if and only if the cross-covariance operator R_{XY} has a decomposition $R_{XY} = R_X^{1/2}TR_Y^{1/2}$, where T is Hilbert–Schmidt and $\| T \| < 1$. When these conditions are satisfied, $I(X, Y) = -\frac{1}{2} \sum_n \log(1 - \gamma_n)$, where $\{\gamma_n, n \geqslant 1\}$ are the non-zero eigenvalues of $P_Y T^* T P_Y$.*

Lemma 3 (Douglas, 1966). *Suppose A and B are two bounded linear operators in$L_2[0, T]$. Then* range$(A) \subset$ range(B) *if and only if there exists a bounded linear operator C such that $A = BC$.* Range(A) = range(B) *if and only if there exists such a C which has a bounded inverse.*

LEMMA 4. *Suppose that* $I(S, N) < \infty$ *and* $I(S, S + N) < \infty$, *where* (S_t) *and* (N_t) *are jointly Gaussian. Then* $\text{range}(R_S^{1/2}) \subset \text{range}(R_N^{1/2})$ *and* $\text{range}(R_{S+N}^{1/2}) = \text{range}(R_N^{1/2})$.

Proof. $R_{SN} = R_S^{1/2} V R_N^{1/2}$, V Hilbert–Schmidt, $\| V \| < 1$, $P_S V = V$, $R_{S,S+N} = R_S + R_{SN} = R_S^{1/2} U R_{S+N}^{1/2}$, U Hilbert–Schmidt, $\| U \| < 1, P_S U = U$. Thus $R_S^{1/2} + V R_N^{1/2} = U R_{S+N}^{1/2}$, so that $R_S + R_{SN} + R_{NS} + R_N^{1/2} V^* V R_N^{1/2} = R_{S+N}^{1/2} U^* U R_{S+N}^{1/2}$. Thus, $R_{S+N} - R_N + R_N^{1/2} V^* V R_N^{1/2} = R_{S+N}^{1/2} U^* U R_{S+N}^{1/2}$, or $R_{S+N}^{1/2}(I - U^*U) R_{S+N}^{1/2} = R_N^{1/2}(I - V^*V) R_N^{1/2}$. The fact that U and V are each compact with norm less than one implies $\text{range}(R_{S+N}^{1/2}) = \text{range}(R_N^{1/2})$. Hence, there exists a bounded linear operator B such that $R_{S+N}^{1/2} = R_N^{1/2} B$. Using this in the above equality $R_S^{1/2} + V R_N^{1/2} = U R_{S+N}^{1/2}$, one has $R_S^{1/2} = -R_N^{1/2} V^* + R_N^{1/2} B U^*$ which implies that $\text{range}(R_S^{1/2}) \subset \text{range}(R_N^{1/2})$, by Lemma 3.

The following theorem is the main result of this paper.

THEOREM 1. *Suppose that* $I(S, N) < \infty$. *Then the following are equivalent*: (a) $I(S, S + N) < \infty$; (b) μ_N *and* μ_{S+N} *are strongly equivalent*; (c) *almost all paths of* (S_t) *belong to* $\text{range}(R_N^{1/2})$.

Proof. Since $I(S, N) < \infty$, $R_{SN}^{1/2} = R_S^{1/2} V R_N$ with V Hilbert–Schmidt, $\| V \| < 1$, and $P_S V = V P_N = V$. Suppose first that $I(S, S + N) < \infty$. Then $R_{S,S+N} = R_S + R_{SN} = R_S^{1/2} U R_{S+N}^{1/2}$ with U Hilbert–Schmidt, $\| U \| < 1$, and $P_S U = U$. Hence $R_S^{1/2} + V R_N^{1/2} = U R_{S+N}^{1/2}$. From Lemma 4, $R_S^{1/2} = R_N^{1/2} P$, P bounded, and so $R_{S+N}^{1/2} = R_N^{1/2} B$, where $BB^* = PP^* + PV + V^*P^* + I$. Hence, $R_S^{1/2} + V R_N^{1/2} = U R_{S+N}^{1/2} = U B^* R_N^{1/2}$, so that $R_S^{1/2} = R_N^{1/2}[V^* + BU^*]$. Since V and U are each Hilbert–Schmidt and B is bounded, we have that $R_S^{1/2} = R_N^{1/2} P$ for P Hilbert–Schmidt. This implies that $\mu_S[\text{range}(R_N^{1/2})] = 1$, so that (a) $\Rightarrow$ (c). Conversely, suppose that $\mu_S[\text{range}(R_N^{1/2})] = 1$, so that $R_S^{1/2} = R_N^{1/2} P$ for P Hilbert–Schmidt. Then $R_{S,S+N} = R_S + R_{SN} = R_S + R_S^{1/2} V R_N^{1/2} = R_S^{1/2} U R_{S+N}^{1/2}$, and $R_{S+N}^{1/2} = R_N^{1/2} B$, B bounded. We show that U is Hilbert–Schmidt and $\| U \| < 1$. Since U can be assumed (Lemma 1) to satisfy $U = U P_{S+N}$, and $\overline{\text{range}(R_{S+N})} \subset \overline{\text{range}(R_N)}$, we assume now that $\overline{\text{range}(R_N)} = L_2[0, T]$ (if this is not satisfied, we can restrict attention to $\overline{\text{range}(R_N)}$). The equalities $R_S^{1/2} = R_N^{1/2} P$ and $R_S + R_S^{1/2} V R_N^{1/2} = R_S^{1/2} U R_{S+N}^{1/2}$ yield $P^* + V = UB^*$, where again $BB^* = PP^* + PV + V^*P^* + I$. Now, the fact that $\| V \| < 1$ implies that B^{-1} exists and is bounded (Baker, 1973a). Since P and V are Hilbert–Schmidt, U must be Hilbert–Schmidt. To see that $\| U \| < 1$, we note that $U^*U = B^{-1}(P^* + V)(P + V^*)B^{*-1} = B^{-1}[BB^* - I + VV^*]B^{*-1} = I + B^{-1}[VV^* - I]B^{*-1}$. Since U is compact, $\| U \| = 1$ only if there exists x in $L_2[0, T]$ such that $U^*Ux = x$, which requires $B^{-1}[VV^* - I]B^{*-1}x = 0$. This cannot be true, since $\| V \| < 1$. Hence (c) $\Rightarrow$ (a), so that (a) $\Leftrightarrow$ (c).

To see that (b) $\Leftrightarrow$ (c), suppose $\mu_{S+N} \sim^s \mu_N$. Then $\text{range}(R_S^{1/2}) \subset \text{range}(R_N^{1/2})$ (Baker, 1977), so that $R_S^{1/2} = R_N^{1/2} P$ for P bounded, and $PP^* + PV + V^*P^*$

must be trace-class. This implies that $PP^* + PV + V^*P^* + V^*V = (P + V^*)(P^* + V)$ is trace-class. Hence, $P + V^*$ must be Hilbert–Schmidt, and since V is Hilbert–Schmidt, P must be Hilbert–Schmidt, and so $\mu_S[\text{range}(R_N^{1/2})] = 1$. Conversely, suppose that $\mu_S[\text{range}(R_N^{1/2})] = 1$. Then $\| V \| < 1$ implies that $\mu_{S+N} \sim \mu_N$ (Baker, 1973a), and we have $R_{S+N} = R_N^{1/2}[I + PV + V^*P^* + PP^*]R_N^{1/2}$. Since P and V are Hilbert–Schmidt, $PV + V^*P^* + PP^*$ must be trace-class. Hence $\mu_{S+N} \sim^s \mu_N$, and thus (b) $\Leftrightarrow$ (c). ∎

It can be seen that Pitcher's result is a special case of Theorem 1, since $I(S, N) = 0$ when S and N are independent ($\mu_{S,N} = \mu_S \otimes \mu_N$); moreover, it is not necessary to assume that $\mu_{S+N} \sim \mu_N$.

One may ask if Theorem 1 can be further improved. The following result shows that substantial improvement is not possible.

THEOREM 2 (Baker, 1978). (1) *There exists a process* (S_t), *with* (S_t) *and* (N_t) *jointly Gaussian, such that* μ_{S+N} *and* μ_N *are orthogonal, while* $I(S, S + N) = 0$ *and almost all paths of* (S_t) *are in range* $(R_N^{1/2})$.

(2) *There exists a process* (S_t), *with* (S_t) *and* (N_t) *jointly Gaussian, such that* μ_{S+N} *and* μ_N *are strongly equivalent and almost all paths of* (S_t) *belong to range* $(R_N^{1/2})$, *but* $I(S, S + N) = \infty$.

(3) *There exists a process* (S_t), *with* (S_t) *and* (N_t) *jointly Gaussian, such that* $I(S, S + N) = 0$, *almost all paths of* (S_t) *lie outside* range$(R_N^{1/2})$, *and* μ_{S+N} *and* μ_N *are orthogonal.*

The statement of Theorem 2 should be interpreted as follows. For any given Gaussian process (N_t) with induced measure μ_N, there exists a Gaussian process (N_t'), defined on a probability space $(\mathscr{A}, \mathscr{F}, Q)$ such that (N_t') has almost all paths (dQ) in $L_2[0, T]$, and the path map of (N_t') induces the Gaussian measure μ_N on the Borel σ-field of $L_2[0, T]$. Moreover, there exists (for any one of (1), (2), (3) of Theorem 2) a process (S_t') on $(\mathscr{A}, \mathscr{F}, Q)$ such that (S_t', N_t') induces a joint Gaussian measure on the Borel σ-field of $L_2[0, T] \times L_2[0, T]$, and this measure has the specified property.

As a consequence of Theorem 2, one sees that the hypothesis $I(S, N) < \infty$ of Theorem 1 cannot be omitted; moreover, it is not possible to exchange this hypothesis with any one of conditions (a), (b), (c) of Theorem 1.

Suppose that $A: L_2[0, T] \to L_2[0, T]$ is a bounded linear operator. It is well-known that $I(AS, AS + N) \leqslant I(S, AS + N)$. The converse does not hold in general if A does not have an inverse. However, the following result can be obtained.

PROPOSITION 1. *If* $I(S, N) < \infty$, *then* $I(AS, AS + N) < \infty$ *implies* $I(S, AS + N) < \infty$.

Proof. Since $I(AS, N) \leqslant I(S, N)$, by Lemma 4 and Lemma 3 there exists a bounded operator B with bounded inverse such that $R_{AS+N}^{1/2} = R_N^{1/2}B$. $R_{SN} = R_S^{1/2}VR_N^{1/2}$ for V Hilbert–Schmidt with $\| V \| < 1$. From Theorem 1, $R_{AS}^{1/2} = R_N^{1/2}P$ for P Hilbert–Schmidt. Hence $R_{S,AS+N} = R_S^{1/2}UR_{AS+N}^{1/2} = R_SA^* + R_{SN} = R_SA^* + R_S^{1/2}VR_N^{1/2}$. Since $\text{range}(U) \subset \overline{\text{range}(R_S)}$ and $\text{range}(V) \subset \overline{\text{range}(R_S)}$, one now has $UR_{AS+N} = UB^*R_N^{1/2} = R_S^{1/2}A^* + VR_N^{1/2} = W^*R_{AS}^{1/2} + VR_N^{1/2}$, where W is a partially isometric operator, isometric on $\overline{\text{range}(R_{AS})}$, satisfying $R_{AS}^{1/2} = AR_S^{1/2}W^*$. Using $R_{AS}^{1/2} = P^*R_N^{1/2}$, the preceding equalities show that $UB^* = W^*P^* + V$ on $\overline{\text{range}(R_N)}$. If $\overline{\text{range}(R_N)}$ is infinite-dimensional, then U is Hilbert–Schmidt because V and P^* are Hilbert–Schmidt, B^* maps $\text{range}(R_N)$ into $\text{range}(R_N)$, and B^{*-1} exists and is bounded. If $\text{range}(R_N)$ is finite-dimensional, then U is Hilbert–Schmidt because U can be taken to satisfy (Lemma 1) $U = UP_{AS+N}$, and $\overline{\text{range}(R_{AS+N})} = \overline{\text{range}(R_N)}$. To show $\| U \| < 1$, and hence $I(S, AS + N) < \infty$, it is sufficient to show that U^*U cannot have $+1$ as an eigenvalue. Now $R_{AS+N} = R_N^{1/2}BB^*R_N^{1/2} = R_{AS} + R_{AS,N} + R_{AS,N}^* + R_N = R_N^{1/2}[PP^* + PWV + V^*W^*P^* + I]R_N^{1/2}$, and thus $BB^* = PP^* + PWV + V^*W^*P^* + I$.

$$
\begin{aligned}
U^*U &= B^{-1}[PWW^*P^* + V^*W^*P^* + PWV + V^*V]B^{*-1} \\
&= B^{-1}[BB^* - I - PP^* + PWW^*P^* + V^*V]B^{*-1} \\
&= I - B^{-1}[V^*V - I + PWW^*P^* - PP^*]B^{*-1}.
\end{aligned}
$$

Thus, $\| U \| = 1$ if and only if $V^*V - I + PWW^*P^* - PP^*$ has zero as an eigenvalue. Since $V^*V \leqslant I$ and $PWW^*P^* \leqslant PP^*$ (because W is partially isometric), $\| U \| = 1$ implies V^*V has $+1$ as an eigenvalue. This cannot occur, since $\| V \| < 1$. ∎

From Theorem 1 and Proposition 1, one obtains the following result.

COROLLARY. *Suppose that* $A: L_2[0, T] \to L_2[0, T]$ *is continuous and linear. If* $I(S, N) < \infty$, *then the following are equivalent*: (a) $I(S, AS + N) < \infty$; (b) $I(AS, AS + N) < \infty$; (c) $AS + N$ *and* N *are strongly equivalent*; (d) *almost all sample paths of* AS *belong to* $\text{range}(R_N^{1/2})$.

The following result can be applied if one is unable to determine whether $I(S, N) < \infty$.

PROPOSITION 2. *Suppose that* $R_{S+N} \geqslant R_N$. *Then* $I(S, S + N) < \infty$ *implies that* μ_N *and* μ_{S+N} *are strongly equivalent, that almost all paths of* (S_t) *belong to* $\text{range}(R_N^{1/2})$, *and that* $I(S, N) < \infty$.

Proof. $R_{S,S+N} = R_S + R_S^{1/2}VR_N^{1/2} = R_S^{1/2}UR_{S+N}^{1/2}$; if $I(S, S + N) < \infty$ then U is Hilbert–Schmidt and $\| U \| < 1$. It is sufficient, by Theorem 1, to show that $R_{S+N} \geqslant R_N$ implies $I(S, N) < \infty$ (V Hilbert–Schmidt and $\| V \| < 1$). We have $R_S^{1/2} + VR_N^{1/2} = UR_{S+N}^{1/2}$, or $R_S + R_{S,N} + R_{N,S} + R_N^{1/2}V^*VR_N^{1/2} = R_{S+N}^{1/2}U^*UR_{S+N}^{1/2} = R_{S+N} - R_N + R_N^{1/2}V^*VR_N^{1/2}$. Thus $R_{S+N}^{1/2}(I - U^*U)R_{S+N}^{1/2} =$

$R_N^{1/2}(I - V^*V)R_N^{1/2} \equiv G$. Since U is Hilbert–Schmidt with $\| U \| < 1$, $\text{range}(R_{S+N}^{1/2}) = \text{range}(G^{1/2})$, from Lemma 3. But $\text{range}(G^{1/2}) \subset \text{range}(R_N^{1/2})$ from Lemma 3 and the definition of G. Hence $R_{S+N}^{1/2} = R_N^{1/2}T$, T bounded. In fact, T^{-1} exists, and is bounded, since $R_{S+N} \geqslant R_N$ implies $\text{range}(R_N^{1/2}) \subset \text{range}(R_{S+N}^{1/2})$. Using $R_{S+N} - R_N + R_N^{1/2}V^*VR_N^{1/2} = R_{S+N}^{1/2}U^*UR_{S+N}^{1/2}$ and $R_{S+N} \geqslant R_N$, we have that $R_{S+N}^{1/2}U^*UR_{S+N}^{1/2} \geqslant R_N^{1/2}V^*VR_N^{1/2}$, or $U^*U \geqslant TV^*VT^*$. This shows that TV^* is Hilbert–Schmidt, since U is Hilbert–Schmidt; since T^{-1} exists and is bounded, V must be Hilbert–Schmidt. The fact that $\| V \| < 1$ follows from $G = R_N^{1/2}(I - V^*V)R_N^{1/2}$, and $\text{range}(G^{1/2}) = \text{range}(R_N^{1/2})$. Thus $I(S, N) < \infty$, and application of Theorem 1 completes the proof. ∎

We next give a necessary condition for $I(S, S + N) < \infty$ and also for μ_{S+N} to be strongly equivalent to μ_N.

Proposition 3. *In order that either* $I(S, S + N) < \infty$ *or* μ_{S+N} *and* μ_N *be strongly equivalent, it is necessary that* $\text{range}(R_S^{1/2})$ *be contained in* $\text{range}(R_N^{1/2})$.

Proof. Suppose it is not true that $\text{range}(R_S^{1/2}) \subset \text{range}(R_N^{1/2})$. Then $\mu_{S+N} \perp \mu_N$ (Baker, 1977). Defining $R_{S,N} = R_S^{1/2}VR_N^{1/2}$ and $R_{S,S+N} = R_S^{1/2}UR_{S+N}^{1/2} = R_S + R_{SN}$, we have $UR_{S+N}^{1/2} = R_S^{1/2} + VR_N^{1/2}$. If $\text{range}(R_{S+N}^{1/2}) \subset \text{range}(R_N^{1/2})$ then $R_{S+N}^{1/2} = R_N^{1/2}G$ for G bounded, giving $R_S^{1/2} = R_N^{1/2}(-V^* + G^*U^*)$. This is a contradiction, by Lemma 3, since $\text{range}(R_S^{1/2})$ is not contained in $\text{range}(R_N^{1/2})$. Hence $\text{range}(R_{S+N}^{1/2})$ is not contained in $\text{range}(R_N^{1/2})$. Using $UR_{S+N}^{1/2} = R_S^{1/2} + VR_N^{1/2}$, one obtains $R_{S+N}^{1/2}U^*UR_{S+N}^{1/2} = R_S + R_{SN} + R_{NS} + R_NV^*VR_N^{1/2} = R_{S+N} - R_N + R_N^{1/2}V^*VR_N^{1/2}$, so that $R_{S+N}^{1/2}(I - U^*U)R_{S+N}^{1/2} = R_N^{1/2}(I - V^*V)R_N^{1/2}$. Now suppose that $I(S, S + N) < \infty$. Then U is Hilbert–Schmidt and $\| U \| < 1$, so that $\text{range}(R_{S+N}^{1/2}) = \text{range}(T^{1/2})$ where $T = R_{S+N}^{1/2}(I - U^*U)R_{S+N}^{1/2}$. However, $\text{range}(T^{1/2}) \subset \text{range}(R_N^{1/2})$, from $T = R_N^{1/2}(I - V^*V)R_N^{1/2}$ and Lemma 3. This implies $\text{range}(R_{S+N}^{1/2}) \subset \text{range}(R_N^{1/2})$, a contradiction. Hence $I(S, S + N) = \infty$. ∎

From Theorem 2, one could have $R_S^{1/2} = R_N^{1/2}P$, with P bounded but not Hilbert–Schmidt, and still have μ_{S+N} strongly equivalent to μ_N and $I(S, S + N) = 0$. The requirement that P be Hilbert–Schmidt is equivalent to requiring that almost all paths of (S_t) belong to $\text{range}(R_N^{1/2})$. If it is not true that $R_S = R_N^{1/2}P$ for some bounded P, then one has that $\text{range}(R_S^{1/2})$ is not contained in $\text{range}(R_N^{1/2})$ (Douglas, 1966). Proposition 3 can be applied without knowing the nature of any statistical dependence between (S_t) and (N_t). As an example, $I(S, S + N) = \infty$ when (S_t) and (N_t) are defined as in any of the following:

(a) (S_t) has covariance $e^{-\alpha|t-s|}$ $(\alpha > 0)$ and (N_t) has covariance $\min(t, s)$;

(b) (S_t) has covariance $e^{-\alpha|t-s|}$ and (N_t) has covariance $T - \max(t, s)$;

(c) (S_t) has covariance $T - \max(t, s)$ and (N_t) has covariance $\min(t, s)$.

Both (a) and (b) continue to hold if $e^{-\alpha|t-s|}$ is replaced by $T - | t - s |$.

These statements can be proved from known results (Baker 1973a). For example if (N_t) has covariance $\min(t, s)$ then the range space of $R_N^{1/2}$ consists of all elements of $L_2[0, T]$ that are equal (a.e. dt) to an absolutely continuous function that vanishes at zero and has $L_2[0, T]$ derivative. If (N_t) has covariance T-$\max(t, s)$, then the range space of $R_N^{1/2}$ consists of all elements which are equal (a.e. dt) to an absolutely continuous function that vanishes at T and has $L_2[0, T]$ derivative. If (S_t) has covariance $e^{-\alpha|t-s|}$, then range$(R_S^{1/2})$ consists of all elements in $L_2[0, T]$ that are equal (a.e. dt) to an absolutely continuous function with derivative in $L_2[0, T]$. The range space is the same if (S_t) has covariance $T - |t - s|$.

Our final result involves stationary processes.

PROPOSITION 4. *Suppose* $(S_t + N_t)$ *are stationary with spectral densities* Φ_{S+N} *and* Φ_N, *with* Φ_N *rational. If* $\Phi_N(\lambda) \leqslant \Phi_{S+N}(\lambda)$ *a.e.* $d\lambda$, *and*

$$\int \{[\Phi_{S+N}(\lambda) - \Phi_N(\lambda)]/\Phi_N(\lambda)\}\, d\lambda < \infty,$$

then $S + N$ *and* N *are strongly equivalent.*

Proof. Pinsker has shown (1964) that if S and N are independent, S has spectral density Φ_S, and N has rational spectral density Φ_N, then $I(S, S + N) < \infty$ if $\int_{-\infty}^{\infty} [\Phi_S(\lambda)/\Phi_N(\lambda)]\, d\lambda < \infty$. From Theorem 1, this also implies $R_S = R_N^{1/2} G R_N^{1/2}$ for G trace-class. $\Phi_{S+N} - \Phi_N$ is a spectral density, and by Pinsker's result, $R_{S+N} - R_N = R_N^{1/2} T R_N^{1/2}$ for T trace-class. Hence $R_{S+N} = R_N^{1/2}(I + T)R_N^{1/2}$. The condition $\Phi_{S+N}(\lambda) \geqslant \Phi_N(\lambda)$ a.e. $d\lambda$ implies range$(R_N^{1/2}) \subset$ range$(R_{S+N}^{1/2})$ (Baker, 1973a); thus $I + T$ must have bounded inverse (Lemma 3). This shows that μ_{S+N} and μ_N re strongly equivalent.

RECEIVED: December 16, 1977; REVISED: October 19, 1978

REFERENCES

C. R. BAKER (1970), Mutual information for Gaussian processes, *SIAM J. Appl. Math.* **19**, 451–458.

C. R. BAKER (1973a), On equivalence of probability measures, *Ann. Probability* **1**, 690–698.

C. R. BAKER (1973b), Joint measures and cross-covariance operators, *Trans. Amer. Math. Soc.* **186**, 273–289.

C. R. BAKER (1977), Absolute continuity for a class of probability measures, *in* "Lecture Notes in Mathematics (Vector Space Measures)" No. 644, pp. 44–450, Springer-Verlag (Berlin).

C. R. BAKER (1978), Relations between mutual information, strong equivalence, signal-to-noise ratio, and signal sample path properties, *in* "Trans. 8th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes," Vol. A, 85–92.

R. L. Dobrushin (1963), General formulation of Shannon's basic theorems on the theory of information, *Amer. Math. Soc. Trans. (2)* **33**, 323–348.

R. G. Douglas (1966), On majorization, factorization, and range inclusion of operators in Hilbert space, *Proc. Amer. Math. Soc.* **17**, 413–415.

J. Feldman (1958), Equivalence and Perpendicularity of Gaussian Processes, *Pacific J. Math.* **8**, 699–708.

J. Hajek (1958), On a property of normal distributions of any stochastic process, *Czechoslovak Math. J.* **8**, 610–618.

J. Hajek (1962), On linear statistical problems in stochastic processes, *Czechoslovak Math. J.* **87**, 404–444.

D. B. Osteyee and I. J. Good (1974), Information, weight of evidence, the singularity between probability measures and signal detection, *in* "Lecture Notes in Mathematics," No. 376, Springer-Verlag (Berlin).

M. S. Pinsker (1964), "Information and Information Stability of Random Variables and Processes," Holden–Day, San Francisco.

T. S. Pitcher (1963), On the sample functions of processes which can be added to a Gaussian process, *Ann. Math. Stat.* **34**, 329–333.

C. R. Rao and V. S. Varadarajan (1963), Discrimination of Gaussian processes, *Sankhyā Ser. A* **25**, 303–330.